

## Diversity in the CDR3 Region of $V_H$ Is Sufficient for Most Antibody Specificities

John L. Xu\* and Mark M. Davis\*†

\*Department of Microbiology and Immunology

†Howard Hughes Medical Institute  
Stanford University School of Medicine  
Stanford, California 94305

### Summary

All rearranging antigen receptor genes have one or two highly diverse complementarity determining regions (CDRs) among the six that typically form the ligand binding surface. We report here that, in the case of antibodies, diversity at one of these regions, CDR3 of the  $V_H$  domain, is sufficient to permit otherwise identical IgM molecules to distinguish between a variety of hapten and protein antigens. Furthermore, we find that somatic mutation can allow such antibodies to achieve surprisingly high affinities. These results are consistent with a model in which the highly diverse CDR3 loops are the key determinant of specificity in antigen recognition in both T cell receptors (TCR) and antibodies, whereas the germline-encoded CDR1 and CDR2 sequences are much more cross-reactive.

### Introduction

During lymphocyte development, a large repertoire of heterodimeric antigen receptors, both antibodies and TCRs, are generated by a variety of mechanisms (Tonegawa, 1983; Davis and Bjorkman, 1988). In antibodies, the binding site for antigens is formed by six CDRs that loop out from the V region backbone formed by two sheets of  $\beta$ -pleated strands (reviewed in Davies et al., 1990). Great importance has been attached to the germline V gene repertoire for the development of effective immune responses, as most of the CDRs are encoded by the germline sequences (with the exception of CDR3 of the heavy chain). It has been postulated that the V region genes are selectively retained in the germline during evolution because of their capacity to accommodate different antigens, especially pathogens (e.g., Cohn et al., 1980; Rajewsky et al., 1987).

Recent structural analyses show that in  $\alpha\beta$  TCRs, amino acids at positions equivalent to the CDRs in antibodies also form the principal contacts between TCRs and their peptide/MHC ligands (Garboczi et al., 1996; Garcia et al., 1996; Ding et al., 1998; Reinherz et al., 1999). While no complete structure of a  $\gamma\delta$  TCR in complex with its ligand is currently available, a number of studies indicate that  $\gamma\delta$  TCRs recognize antigens in an antibody-like manner (reviewed in Chien et al., 1996; see also Li et al., 1998). The binding interfaces between antigen receptor molecules and their ligands are generally large (over 1500 Å<sup>2</sup>) and 10–30 side chains from each side make close contacts (reviewed in Davies et al.,

1990; Davies and Cohen, 1996; Garcia et al., 1999). However, despite the broad interface seen in the crystal structures, sequence diversity in antigen receptors is not evenly distributed among all six CDRs but is highly concentrated in one (in the cases of antibodies and  $\gamma\delta$  TCRs) or two (in the case of  $\alpha\beta$  TCRs) CDRs (Davis and Bjorkman, 1988; Davis and Chien, 1999). Although the skewing of diversity toward CDR3s in  $\alpha\beta$  TCRs is understandable because these amino acid residues mainly recognize the antigenic peptide while other CDRs primarily interact with MHC (Jorgensen et al., 1992; Garboczi et al., 1996; Garcia et al., 1996; Sant'Angelo et al., 1996; Ding et al., 1998; Reinherz et al., 1999), there has been no explanation for such a phenomenon in antibodies and  $\gamma\delta$  TCRs. Even more puzzling is the finding that in antibodies, both CDR1 and CDR2 of the heavy chains and all CDRs of the light chains have only a few "canonical" conformations, and only the heavy chain CDR3 loop has a wide range of variations in both length and shape (Chothia and Lesk, 1987; Chothia et al., 1989). Such canonical CDR loop shapes have also been seen in  $\alpha\beta$  TCR crystal structures (CDR2 of  $\alpha$  and  $\beta$  particularly, but not in CDR3 $\alpha$  or CDR3 $\beta$ ) (Garcia et al., 1999), although some of this uniformity may relate to MHC binding requirements. The skewing of diversity in all antigen receptor molecules has led to the suggestion that the highly diverse CDR3 sequences are the principal determinant of specificity in antigen recognition, at least in the primary repertoire (Davis et al., 1997, 1998).

To further understand the molecular basis of antigen receptor specificity, we have constrained mice to use a single  $V_H$  gene but full CDR3 diversity to generate their B cell repertoire. We challenged such mice with a variety of protein and hapten antigens and monitored the development of primary and memory immune responses. We find that antigen-specific IgM molecules isolated from primary response of these mice can differ only in the CDR3 of the  $V_H$  domain; furthermore, we find that somatic mutation can allow such antibodies to achieve surprisingly high affinities upon rechallenge with protein antigens. The only antigens that seem generally unable to be accommodated by an arbitrarily chosen  $V_H$  are bacterial polysaccharides. These results indicate that an extensive  $V_H$  (both  $V_H$  and  $V_L$ ) gene repertoire is not necessary for the production of specific antibodies to most antigens, and that the  $V_H$  CDR3 plays a very different role in the makeup of the antibody binding site than the germline-encoded CDR1 and CDR2 sequences. These results suggest that the purpose of the highly diverse CDR3 region of all antigen receptors, both TCRs and antibodies, is to determine antigen specificity.

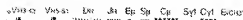
### Results

#### Experimental System

In order to characterize the antibody responses of an organism with a very limited V region repertoire but full

†To whom correspondence should be addressed (e-mail: mdavis@cmgm.stanford.edu).

# A. Heavy Chain Minilocus, HC1



# B. $\lambda$ Light Chain Locus

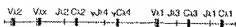


Figure 1. Physical Map of the Transgenic Heavy Chain Minilocus *HC1* and the Endogenous Mouse  $\lambda$  Light Chain Locus. Drawings are schematic and not to scale.  $\delta$  denotes pseudogenes. V, variable gene segment; D, diversity segment; J, joining segment; C, constant region gene; S, switch region; E, enhancer.

CDR3 diversity, we utilize a previously described transgene, *HC1*, which contains one functional human  $V_H$  gene segment ( $V_H5-51$ ), ten functional D segments, six  $J_H$  segments, and the  $C_H$  and  $C_H1$  constant regions (Figure 1A; Taylor et al., 1992, 1994). The *HC1* transgenic

mice were bred onto an *IgH<sup>-/-</sup> IgK<sup>-/-</sup>* background (Chen et al., 1993a, 1993b). With this genotype (*HC1<sup>tg</sup> IgH<sup>-/-</sup> IgK<sup>-/-</sup>*), the only functional immunoglobulin that can be produced would combine the single  $V_H5-51$  with one of the three  $\lambda$  light chain V regions ( $V_L1$ ,  $V_L2$ , or  $V_L3$ ; Figure 1B) from the mouse. Flow cytometry analysis revealed that the *HC1* transgene partially rescues the B cell compartment of an *IgH<sup>-/-</sup> IgK<sup>-/-</sup>* mouse, although B cell numbers are extremely low (5%–10% normal; data not shown). Serum analysis shows that levels of IgM and IgG1 range from 60 to 300  $\mu$ g/ml (5%–10% normal) and 0.3 to 5.0  $\mu$ g/ml (0.1%–1% normal), respectively, in these mice (data not shown).

## Antigen-Specific Development of Immune Responses in the *HC1<sup>tg</sup> IgH<sup>-/-</sup> IgK<sup>-/-</sup>* Mice

To test whether the B cell repertoire with diversity only in  $V_H$  CDR3 is effective, we immunized the *HC1<sup>tg</sup> IgH<sup>-/-</sup> IgK<sup>-/-</sup>* mice with a variety of protein antigens: keyhole limpet hemocyanin (KLH), hen egg-white lysozyme (HEL),

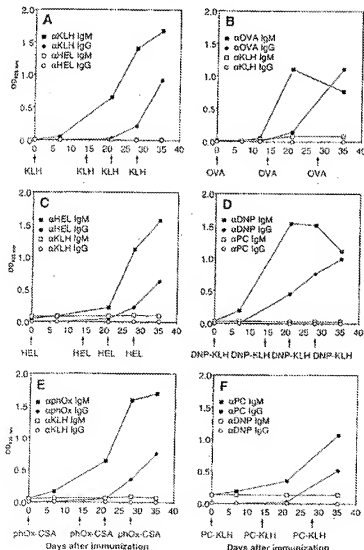


Figure 2. Antigen-Specific Development of IgM and IgG Responses in the *HC1<sup>tg</sup> IgH<sup>-/-</sup> IgK<sup>-/-</sup>* Mice

The *HC1<sup>tg</sup> IgH<sup>-/-</sup> IgK<sup>-/-</sup>* mice were immunized with KLH (A), OVA (B), HEL (C), DNP-KLH (D), phOx-CSA (E), or PC-KLH (F) on indicated days. Serum samples were collected 7 days after each injection and diluted 1:10–1:250 into microtiter wells coated with KLH or HEL (A), OVA or KLH (B), HEL or KLH (C), DNP-BSA or PC-BSA (D), phOx-BSA or KLH (E), or PC-BSA or DNP-BSA (F). Antigen-bound IgM or IgG was detected as described in the Experimental Procedures. At least three mice were tested for each antigen and for clarity, results from a representative mouse were shown. Variations among individual mice were within a 5-fold range.



Figure 3. Sequence Analysis of Monoclonal Anti-DNP, Anti-OVA, Anti-KLH, and Anti-HEL IgM Antibodies Originated from the  $HC1^{-/-}$  IgM $^{+}$  IgG $^{-}$  Mice

The cDNA encoding monoclonal antibodies against various antigens was synthesized and amplified by RT-PCR, cloned, and sequenced. Shown are the deduced amino acid sequences of both heavy (A) and light (B) chains of various monoclonal antibodies. The top line in (A) is the germline V<sub>H</sub>-51 sequence, the top two lines in (B) are the germline V<sub>L</sub>-1 and V<sub>L</sub>-2 sequences. Dashes in the sequences indicate identity to the top line. The CDRs in both heavy and light chain are underlined. In V<sub>H</sub> CDR3, amino acid residues encoded by the J segment are separated by spaces from those encoded by the D segment (in both and N nucleotides). D segment assignments are based on at least six nucleotides of homology; amino acid residues not accounted for by D are presumably encoded by the N nucleotides. All monoclonal antibodies shown here are of IgM isotype. Some of these hybridomas (e.g., DNP-8, DNP-11, KLH-1, KLH-3, and OVA-2) were isolated twice from the same fusion.

chicken ovalbumin (OVA), cholera toxin B subunit (CTB), as well as haptens: 2,4-dinitrophenyl (DNP, conjugated to KLH), 2-phenyl-4-ethoxymethyl-5-oxazolone (phOx, conjugated to chicken serum albumin, CSA), and phosphorylcholine (PC, conjugated to KLH). Figure 2 shows the results of enzyme-linked immunosorbent assay (ELISA) of some of these responses. As typical of normal immune responses, IgM antibodies appeared first, followed by IgG. All  $HC1^{-/-}$  mice examined responded to all antigens tested, except the pigeon cytochrome c (PCC) protein, to which only three out of nine mice were able to make a detectable response (data not shown), suggesting that B cell clones with this specificity are very rare in these mice.

#### Sequence Analysis of Antigen-Specific Antibodies Derived from the $HC1^{-/-}$ IgH $^{+}$ IgG $^{-}$ Mice

It has been previously reported that in mice with a fixed VDJ<sub>H</sub> rearrangement, V gene replacement and somatic hypermutation can generate a significant level of functional diversity in the primary repertoire (Cascaho et al., 1998; López-Macías et al., 1999). It was therefore possible that secondary rearrangements and/or somatic mutations could contribute to the antibody responses seen in the  $HC1^{-/-}$  IgH $^{+}$  IgG $^{-}$  mice. To control for this, as well as to characterize the responding antibodies in

more detail, we immunized  $HC1^{-/-}$  IgH $^{+}$  IgG $^{-}$  mice with antigens and isolated hybridoma cell lines producing antigen-specific IgM antibodies. Animals were immunized only once to minimize somatic hypermutation. Complementary DNA (cDNA) from both heavy and light chain mRNA of these antibodies was synthesized, amplified by polymerase chain reaction (PCR), and then cloned and sequenced. All the hybridomas characterized used both V<sub>H</sub>-51 and C<sub>H</sub> from the  $HC1$  transgene in their heavy chain and used either V<sub>L</sub>-1-J<sub>L</sub> or V<sub>L</sub>-2-J<sub>L</sub> in their light chain (with the exception of hybridoma HEL-2, which appears to use V<sub>L</sub>-1-J<sub>L</sub> in its light chain; Figure 3 and data not shown). No V<sub>H</sub>x usage was observed.

A comparison of the anti-DNP, anti-KLH, anti-OVA, and anti-HEL IgM sequences shows that most express the V<sub>H</sub>-51 gene segment in its germline form and thus differ only in the CDR3 region of the heavy chain. In addition, there are two large sets of hybridomas that express identical light chains as well. The first set includes DNP-6, DNP-7, DNP-8, DNP-11, DNP-12, KLH-5, OVA-1, OVA-2, and OVA-4; and the second includes DNP-10, KLH-6, OVA-6, and HEL-1 (Figure 3). ELISA assays on all of these antibodies confirmed that they only bind the antigen they were raised against, consistent with the whole serum analysis shown in Figure 2

(A) HEL was passed over the immobilized monoclonal HEL-3 antibody (250 RIU) at concentration of 2.5 (triangles), 5.0 (diamonds), 10 (squares), and 20 (circles) nM at a flow rate of 100  $\mu$ l/min.

(B) OVA-7 (20  $\mu$ g/ml) was passed over immobilized CVA (1000 RIU, closed circles) or empty surface (open circles).

(C-D) HCl + light or C<sub>18</sub> mice were immunized with HEL-3 or OVA (1  $\mu$ g). Antisera samples were collected after fourth, fifth, and sixth injection and pooled. Total IgG1 was purified from preimmune (open circles) or antiserum (closed circles) with affinity columns of goat anti-mouse IgG1 (IgG1 specific) antibody coupled to agarose beads, and passed over HEL-3 (C) or CVA (D) coupled to the CMS sensor chip at a concentration of 25  $\mu$ g/ml.

possible, we did find that OVA-7 dissociates very slowly: the apparent dissociation rate is  $\sim 5.0 \times 10^{-4} \text{ s}^{-1}$ . As bivalent antibodies have dissociation rates that are approximately 10-fold slower than their monovalent Fab fragments (Mason and Williams, 1986), the estimated bivalent dissociation rate constant of OVA-7 is approximately  $5.0 \times 10^{-3} \text{ s}^{-1}$ . Given a typical association rate of  $0.5^{-5} \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$  for protein-protein interactions (Northrup and Erickson, 1992), the  $K_D$  of OVA-7 would be in the 1–10 nM range, comparable with monoclonal antibodies isolated from normal mice.

Total IgG1 antibodies from the  $HC1^{+/+} IgH^{-/-} Ig\kappa^{-/-}$  mice immunized with HEL, OVA, or KLH were also used in SPR analysis (Figures 5C and 5D and data not shown). Again, we found that polyclonal IgG1 antibodies purified from the  $HC1^{+/+} IgH^{-/-} Ig\kappa^{-/-}$  mice bind their cognate antigen fairly tightly, with an apparent dissociation rate of  $\sim 1 \times 10^{-4} \text{ s}^{-1}$  (Figures 5C and 5D). Thus, the average  $K_D$  of these antibodies would also be in the 0.2–2 nM range, assuming a typical association rate of  $0.5^{-5} \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ .

Taken together, the antigen binding data from both HEL-3 and OVA-7 monoclonals and the polyclonal antisera shows that somatic mutation can allow antibodies from the  $HC1^{+/+} IgH^{-/-} Ig\kappa^{-/-}$  mice to achieve very high affinities against at least protein antigens, comparable to mice with a full V region repertoire.

#### Failure of the $HC1^{+/+} IgH^{-/-} Ig\kappa^{-/-}$ Mice to Respond to Bacterial Polysaccharides

If one  $V_H$  gene is able to accommodate most or all of the antigens, why are there so many different  $V_H$  genes in most vertebrate genomes? It is currently thought that the T cell-independent IgM response to bacterial polysaccharides provides an important protection to the host during infections (Janeway and Travers, 1997). If this defense relies on a large V region repertoire, as previously suggested (Cohn et al., 1980; Rajewsky et al., 1987), the  $HC1^{+/+} IgH^{-/-} Ig\kappa^{-/-}$  mice may not be able to mount an antibody response to some bacterial carbohydrate antigens. Therefore, we injected five  $HC1^{+/+} IgH^{-/-} Ig\kappa^{-/-}$  and five C57BL/6 (B6) mice with dextran B312 (DEX) from *Leuconostoc mesenteroides*, a near-linear glucose polymer with 96% (41–6) linkages, and followed the immune response by ELISA. DEX-specific IgM was detected in every B6 mouse after the first injection (Figure 6B); however, no IgM response was observed in any of the  $HC1^{+/+} IgH^{-/-} Ig\kappa^{-/-}$  mice even after three injections of DEX (Figure 6A). The  $HC1^{+/+} IgH^{-/-} Ig\kappa^{-/-}$  mice also failed to respond to levan, a branched fructose polymer from *Erwinia herbicola* (Figures 6C and 6D). In contrast, the  $HC1^{+/+} IgH^{-/-} Ig\kappa^{-/-}$  (as well as B6) mice responded to other T-independent antigen, for example, 2,4,6-trinitrophenyl (TNP)-Ficoll (Figures 6E and 6F). Thus, the lack of responsiveness to the two polysaccharide antigens in the  $HC1^{+/+} IgH^{-/-} Ig\kappa^{-/-}$  mice is probably not due to lack of T cell help. Since all anti-DEX monoclonal antibodies reported use the  $\kappa$  light chain (Akolkar et al., 1987; Wang et al., 1990), and it has been shown that the  $V_K$  domains do provide antigenic contacts in these antibodies (Wallick et al., 1989), it is possible that the failure of the  $HC1^{+/+} IgH^{-/-} Ig\kappa^{-/-}$  mice to respond to DEX could be solely due to the lack of the appropriate

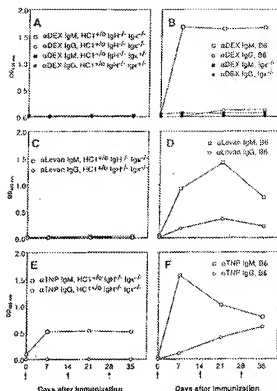


Figure 6.  $HC1^{+/+} IgH^{-/-} Ig\kappa^{-/-}$  Mice Mount IgM Response to TNP-Ficoll but Not to DEX or Levan

Mice with indicated genotypes were injected intravenously with 10  $\mu$ g DEX (A and B) or TNP-Ficoll (E and F), or intraperitoneally with 50  $\mu$ g levan (C and D) on day 0, 14, and 28. Serum samples were collected 7 days after each injection and diluted 1:10–1:250 into microtiter wells coated with DEX (A and B), levan (C and D), or TNP-BSA (E and F). Antigen-binding human or mouse antibodies (IgM and IgG) were detected with alkaline phosphatase conjugated goat anti-human or goat anti-mouse secondary antibodies.

$V_K$  chain(s). To test this possibility, we reintroduced a functional  $ig\kappa$  locus into the mutant mice, but the resulting animals ( $HC1^{+/+} IgH^{-/-} Ig\kappa^{+/+}$ ) were still unable to respond to DEX (Figure 6A). In addition, we also immunized  $ig\kappa^{-/-}$  mice (with a normal  $V_K$  gene repertoire) with this antigen, and they were also unable to respond (Figure 6B).

The observation that ( $igH^{+/+}$ )  $ig\kappa^{-/-}$ ,  $HC1^{+/+} IgH^{-/-} Ig\kappa^{+/+}$ , and  $HC1^{+/+} IgH^{-/-} Ig\kappa^{-/-}$  mice all failed to respond to DEX indicates that neither  $V_H$  5–61 nor any of the  $V_H$  genes can be utilized to accommodate this antigen. However, it cannot be ruled out that a very low frequency of DEX-specific B cells were still present in these mice, but the few precursor B cells were overwhelmed by an excess of antigen (i.e., high zone tolerance was induced) (Howard et al., 1975; Fernandez and Möller, 1977).

#### Discussion

It has been known for some time that in both B and T cell recognition, some epitopes show restricted V gene usage. For example, very limited  $V_H$  and  $V_L$  genes are

used in immune responses of inbred mice to many haptens, polysaccharide, and even protein antigens, including p-azophenylarsenate (Pawlak et al., 1973; Mäkelä et al., 1976; Manser, 1990), phenylazalone (Mäkelä et al., 1976; Kaartinen et al., 1983; Berek and Milstein, 1987), (4-hydroxy-3-nitrophenyl)acetate (NP) (Imanishi and Mäkelä, 1974, 1975; Berek and Bothwell, 1968), DNP (Dzierzak et al., 1980, 1985), phosphorylcholine (Potter and Lieberman, 1970; Sher and Cohn, 1972; Barstad et al., 1974; Crews et al., 1981), phosphatidylcholine (Seidl et al., 1997), dextran (Blomberg et al., 1972; Riblet et al., 1975; Wang et al., 1990), galactan (Mushinski and Potter, 1978), Staphylococcal nuclease (Fathman et al., 1977), and influenza hemagglutinin (McKean et al., 1984; Clarke et al., 1985). T cell recognition of peptide/MHC complexes is often (but not always) fairly restricted with respect to V gene usage (e.g., Hedrick et al., 1988). However, it is not clear whether the preferred V gene segment(s) are absolutely required for such epitopes or are merely the result of extensive clonal expansion where a slight advantage in binding would allow the most optimal  $V_H/V_L$  combination to predominate. In the current study, we find that mice having only one functional  $V_H$  and (effectively) two  $V_L$  gene segments are able to mount highly specific antibody responses to most antigens. It thus seems likely that the restricted V gene usage observed in most immune responses represents the consequence of antigen-driven clonal expansion of kinetically favorable antigen receptor clones(s).

In the  $HCT^{198} IgH^{+/-} Igk^{+/-}$  mice, we find a remarkable restriction in  $V_H$  CDR3 diversity with respect to both length and amino acid composition in the primary immune response to both hapten and protein antigens. In the IgM antibodies, there are many identical or very similar sequences. A restrictive pattern of  $V_H$  CDR3 would normally be interpreted to mean antigen-driven clonal expansion; however, considering the low efficiency of cell fusion, the isolation of identical repeats suggests that certain antibody specificities could be specifically expanded even prior to immunization (Rajewsky et al., 1987). It is not clear at present whether the expansion of a particular specificity occurs as a result of preferred V(D)J rearrangement event or is the result of some type of selection.

Our results also indicate that the antigen binding site of nonsonically mutated antibodies is not uniformly specific but instead consists of one highly diverse CDR loop ( $V_H$  CDR3) and at least four "generic" antigen binding regions (CDRs 1 + 2). Even the somewhat diverse  $V_L$  CDR3 region seems to play a very minor role. Using a phage display library, Winter and colleagues previously showed that a range of hapten and protein binding activities could be isolated from a repertoire of antibodies comprising 50 human  $V_H$  gene segments in combination with a fixed light chain (Hoogenboom and Winter, 1992; Nissim et al., 1994). That  $V_L$  CDR3 might play a dominant role in antigen binding was also suggested by earlier studies showing that a heavy chain alone or even single  $V_H$  domains can bind antigens with a comparable affinity as the intact antibody (Ward et al., 1989; Noel et al., 1996). In addition, it has been found that randomly mutagenizing the  $V_H$  CDR3 of an anti-tetanus toxin antibody enables the isolation of a new specificity (i.e., for fluorocoin) (Barbas et al., 1992). Our results are also consistent

with the findings that murine B cells that have a rearranged heavy chain immunoglobulin gene in the germline always change the CDR3 sequence and often the  $V_H$  as well when they deviate from the original antigen reactivity (Cascaho et al., 1996; López-Macias et al., 1999). Our data may explain the long-standing puzzle of "canonical" CDR structures in antibodies (Chothia and Lesk, 1987; Chothia et al., 1989). We would suggest that uniform CDR shapes are important for the stability of a binding site, with the exception of  $V_H$  CDR3, which has to "fit" the antigen surface much more precisely. These results are also reminiscent of the case of human growth hormone and its receptor, where structural analysis shows that about 30 side chains from each protein are in close contact but only a fraction of those residues have any contribution to the binding affinity and only two account for more than three-quarters of the free energy (Wells, 1996).

Another interesting issue raised by the data presented here is the role of somatic mutation in producing higher-affinity antibodies upon rechallenge with antigen. Our data shows that, at least for protein antigens, having an arbitrarily chosen human  $V_H$  and one of two mouse  $V_L$  chains as a germline repertoire is no barrier to generating very high-affinity IgG antibodies. This finding is consistent with work on the chicken immunoglobulin genes, where only a single functional  $V_H$  and  $V_L$  are utilized, but goes beyond this in two important respects. One is that chickens have somatically mutated antibodies prior to immunization (Weill and Reynaud, 1967; McCormack et al., 1991), thus allowing the argument that this sequence diversification takes the place of a large V region repertoire. Second, the germline  $V_H$  and  $V_L$  genes utilized have been selected over many millions of years for their compatibility with many different antigens and thus they may represent the most polyfunctional  $V_H/V_L$  pair. In our system, we have chosen our  $V_H/V_L$  pair at random, suggesting that all or most such pairs are capable of binding a large number of different antigens, given a single, diverse  $V_H$  CDR3. This indicates that an antibody binding site excluding  $V_H$  CDR3 is exceptionally cross-reactive, at least until acted on by somatic mutation (Patten et al., 1996; Wedemeyer et al., 1997). This interpretation is consistent with the observations that the residues characteristic of antibody binding sites have an unusually high proportion of asparagines, lysines, and aromatic amino acids, more typical of the interior of a globular protein than to its surface and capable of making a wide variety of different contacts (Kabat et al., 1977; Janin and Chothia, 1990; Padian, 1990; Mian et al., 1991). It also has some similarities to the proposal of Pauling and others that a single antibody molecule could fold around different antigens in different ways to achieve specificity (Pauling, 1943).

The data presented here, as well as previous work, strongly suggest that the purpose of highly diverse CDR3 regions in all antigen receptors is to provide antigen specificity. It supports our earlier speculations (Davis et al., 1997, 1998) but is not consistent with the Protection model of Cohn and colleagues, which argues that the primary antibody repertoire is solely defined by germline  $V_H$  and  $V_L$  gene sequences (i.e., CDR1 and CDR2) and that CDR3 is not important in the initial determination of antibody specificity (Langman and Cohn,

The only deficit we have detected in the immune repertoire of mice with limited V gene(s) so far is in their inability to respond to bacterial polysaccharide antigens. Interestingly, chickens are unable to mount robust antibody responses to carbohydrates as well [Jourdain et al., 1998], although this point is controversial [Granfors et al., 1982; Jaakkonen et al., 1983]. If the failure to respond to bacterial polysaccharides is due to the lack of appropriate V<sub>H</sub>/V<sub>L</sub> pair(s) to accommodate these molecules, the HCT-1<sup>-/-</sup>IghT<sup>+/-</sup>Igκ<sup>+/+</sup> mice may have difficulties in clearing bacterial infections. The system described here would provide an experimental test for the idea that unique V<sub>H</sub>/V<sub>L</sub> pairs could be selected by pathogens during evolution and then fixed in the germline [Langman and Cohn, 1987; Cohn and Lanman, 1990].

## Nigv

## Readings

#### Immunization, ELISA, and Hybridomas

diluted serum, and antigen-bound human or mouse IgM or IgG was detected as above with alkaline phosphatase conjugated polyclonal goat anti-human (or mouse) IgM or goat anti-human (or mouse) IgG antibodies. To generate antigen-specific hybridomas, splenocytes were isolated from mice immunized with KLH, OVA, HEL, or UNIP-KLH, and fused with the nonsecreting myeloma P3X63-Ag8.653 cells using standard procedures (Hawley and Lane, 1989). Hybridomas were subcloned by limiting dilution after being shown to produce antigen-specific IgM or IgG antibody in ELISA.

Total RNA was isolated from B cell hybridomas with the RNeasy mini kit (Qiagen). cDNA was synthesized using oligo (dT)<sub>18</sub> (Roche Collaborative Biomedical Products, Bedford, MA) and SUPERSCRIPT RNase H<sup>-</sup> reverse transcriptase (QIAGEN-BRL, Gaithersburg, MD) and amplified by PCR with TaqPlus Precision PCR system (Stratagene, La Jolla, CA). The primer pairs used in PCR were as follows: 5'-51-L1 (5'-octagetcctcctcctcctcctc-3') and C<sub>1</sub> (5'-gaggaagctcctctgagagcagc-3') for  $\mu$  chains; 5'-51-L1 (5'-accatcctcctcctcctcctc-3') and C<sub>1</sub> (5'-gctgagctcctcctcctcctc-3') for  $\gamma$  chains; and V<sub>H</sub>1 (5'-gctgagctcctcctcctcctc-3') and J<sub>H</sub>1.2 (5'-ggagagctcctcctcctcctc-3') for  $\gamma$  chains. Amplified cDNA was cloned into pGEM-T Easy vector (Promega, Madison, WI) and sequenced with the Taq Easy Deoxy Terminator Cycle Sequencing kit (Applied Biosystems, Foster City, CA) on an Applied Biosystems 373A DNA Sequencer (PAN Facility, Stanford University).

IgG1 antibody was purified from mouse serum or hybridoma supernatant by affinity column using goat anti-human IgG1 (Fc specific) antibodies conjugated to agarose beads (Sigma) according to standard procedures (Harlow and Lane, 1988), and exchanged into PBS using Centricon-30 (Amicon, Inc., Beverly, MA). Antigenic Biorisor was immobilized to the Biorace CM5 chip (Pharmacia Biosensor) by standard auro chemistry. Binding of antibody to immobilized antigen was performed in PBS with 20  $\mu$ l injections at a flow rate of 5  $\mu$ l/min. Eluting of HEL freshly purified on a Superdex-75 size exclusion column (Pharmacia) to immobilized antibody was carried out in PBS with 150  $\mu$ l injections of protein at a flow rate of 100  $\mu$ l/min. Regeneration of the biosensor surface was carried out with 100  $\mu$ l of 0.1 M glycine pH 2.0 at a flow rate of 100  $\mu$ l/min. All SPR measurements were performed at 25°C on a Biorace 1000. The association and/or dissociation phases of antigen-antibody interaction were fit with the single-site binding model of BIAevaluation 3.0 (Pharmacia Biosensor).

We thank N. Lomborg, J. D. Capra, and M. Wahl for providing the *HCT* transgenic and *IgH<sup>+</sup> Igε<sup>+</sup>* knockout mice; T. Yao for assistance in mice; J. J. Bonfante for patient instruction on the use of the BiAcore instrument and expert advice on data analysis; N. Lomborg, J. D. Capra, and C. C. Goodnow for helpful discussions; and K. C. Garcia and M. Weigert for comments on the manuscript. J. L. X. was supported by a postdoctoral fellowship from the Damon Runyon-Walter Winchell Foundation (DPG-1412). This work was funded through grants (to M. M. D.) from the National Institute of Health and the Howard Hughes Medical Institute.

## References

- Akolkar, P.N., Sinker, S.K., Bhattacharya, S.B., Liao, J., Grzesio, F., Morrison, S.L., and Kabat, E.A. (1987). Different  $V_H$  and  $V_L$  germ-line genes are used to produce similar combining sites with specificity for  $\alpha$ (1-6) dextrans J. Immunol. 138, 4472-4479.
- Barbas, C.F.D., Bain, J.D., Hoekscha, D.M., and Lerner, R.A. (1992). Semisynthetic combinatorial antibody libraries: a chemical solution to the diversity problem. Proc. Natl. Acad. Sci. USA 89, 4457-4461.
- Barstard, P., Rudloff, S., Potter, M., Cohn, M., Koenigsberg, W., and

- Hood, L. (1974). Immunoglobulin structure: amino terminal sequences of mouse myeloma proteins that bind phosphorylcholine. *Science* **183**, 962-966.
- Burek, C., and Milstein, C. (1987). Mutation drift and repertoire shift in the maturation of the immune response. *Immunol. Rev.* **96**, 23-41.
- Bäer, P.R., and Bothwell, A.L. (1989). The immune response to the hapten NP in C57BL/6 mice: insights into the structure of the B-cell repertoire. *Immunol. Rev.* **105**, 27-43.
- Blomberg, B., Geckler, W.R., and Weigert, M. (1979). Genetics of the antibody response to dextran in mice. *Science* **117**, 176-180.
- Casacub, M., Ma, A., Lee, S., Masal, L., and Wabl, M. (1996). A quasi-monoclonal mouse. *Science* **272**, 1649-1652.
- Chen, J., Trounstein, M., Alt, F.W., Young, F., Kurahara, C., Loring, J.F., and Huszar, D. (1993a). Immunoglobulin gene rearrangement in B cell deficient mice generated by targeted deletion of the  $\lambda$  locus. *Int. Immunol.* **5**, 647-656.
- Chen, J., Trounstein, M., Kurahara, C., Young, F., Kuo, C.C., Xu, Y., Loring, J.F., Alt, F.W., and Huszar, D. (1993b). B cell development in mice that lack one or both immunoglobulin kappa light chain genes. *EMBO J.* **12**, 821-830.
- Chlen, Y.H., Jones, R., and Crowley, M.P. (1995). Recognition by  $\gamma/\delta$  T cells. *Annu. Rev. Immunol.* **14**, 511-532.
- Chothia, C., and Lesk, A.M. (1987). Canonical structures for the hypervariable regions of immunoglobulins. *J. Mol. Biol.* **196**, 901-917.
- Chothia, C., Lesk, A.M., Yemontano, A., Levitt, M., Smith-Gill, S.J., Ar, G., Sheriff, S., Padan, E.A., Davies, D., Tulip, W.R., et al. (1989). Conformations of immunoglobulin hypervariable regions. *Nature* **342**, 871-883.
- Clark, S.H., Huxley, K., Ruzicki, D., Slaut, L., Gornard, W., and Weigert, M. (1985). Inter- and intrachain diversity in the antibody response to influenza haemagglutinin. *J. Exp. Med.* **161**, 687-704.
- Cohn, M., and Langman, R.E. (1990). The protective unit of humoral immunity selected by evolution. *Immunol. Rev.* **115**, 11-47.
- Cohn, M., Langman, R.E., and Geckler, W. (1980). Diversity 1980. In *Immunology 1980: Progress in Immunology IV* (4th Intl. Congr. Immunol.), M. Foixereau and J. Dausset, eds. (New York: Academic Press), pp. 153-201.
- Crowe, S., Orlitz, J., Huang, H., Calame, K., and Hood, L. (1991). A single  $\gamma$  gene segment encodes the immune response to phosphorylcholine: somatic mutation is correlated with the class of the antibody. *Cell* **25**, 59-66.
- Davies, D.R., and Cohen, G.H. (1996). Interactions of protein antigens with antibodies. *Proc. Natl. Acad. Sci. USA* **93**, 7-12.
- Davies, D.R., Padan, E.A., and Sheriff, S. (1990). Antibody-antigen complexes. *Annu. Rev. Biochem.* **59**, 439-473.
- Davis, M.M., and Bjorkman, P.J. (1988). T-cell antigen receptor genes and T-cell recognition. *Nature* **334**, 395-402.
- Davis, M.M., and Chen, Y.-H. (1990). T-cell antigen receptors. In *Fundamental Immunology*, Fourth Edition, W.E. Paul, ed. (Philadelphia, PA: Lippincott-Raven Publications), pp. 341-365.
- Davis, M.M., Lyons, D.S., Altman, J.D., McHeyzer-Williams, M., Hamel, J., Bonifaz, J., and Chen, Y. (1997). T cell receptor biochemistry, repertoire selection and general features of TCR and Ig structure. *Ciba Found. Symp.* **204**, 94-100 (Discussion, 100-104).
- Davis, M.M., Bonifaz, J.J., Rensch, Z., Lyons, D., Hamel, J., Arden, S., and Chen, Y. (1998). Ligand recognition by  $\alpha\beta$  T cell receptors. *Annu. Rev. Immunol.* **16**, 523-544.
- Ding, Y.-H., Smith, K.J., Gortoczi, D.N., Utz, U., Biddison, W.E., and Wiley, D.C. (1993). Two human T cell receptors bind in a similar diagonal mode to the HLA-A2/Tax peptide complex using different TCR amino acids. *Immunity* **8**, 403-411.
- Dzierzak, E.A., Janeway, C.A., Jr., Rosenstein, R.W., and Goffb, P.D. (1989). Expression of an  $\alpha$ 1 (p6-400) domain in vivo anti-dendrocyte antibody responses. I. Mapping of genes for Id-400 expression to the variable region of immunoglobulin heavy-chain locus and to the variable region of immunoglobulin  $\kappa$ -light-chain locus. *J. Exp. Med.* **152**, 720-729.
- Dzierzak, E.A., Brodeur, P., Marion, T., Janeway, C.A., Jr., and Bothwell, A. (1985). Molecular characterization of antibodies bearing Id-400. II. Molecular basis for Id-400 expression. *J. Exp. Med.* **162**, 1494-1511.
- Fathman, C.G., Piselsky, D.S., and Sachs, D.H. (1977). Genetic control of the immune response to nucleases. V. Genetic linkage and strain distribution of anti-nuclease idiotypes. *J. Exp. Med.* **145**, 569-577.
- Fernandez, C., and Möller, G. (1977). Immune response against two epitopes on the same thymus-independent polysaccharide carrier. I. Role of epitope density in carrier-dependent immunity and tolerance. *Immunology* **33**, 59-68.
- Garbocci, D.N., Ghosh, P., Utz, U., Fan, C.H., Biddison, W.E., and Wiley, D.C. (1996). Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* **384**, 134-141.
- Garcia, K.C., Degano, M., Stanfield, R.L., Brumback, A., Jackson, M.R., Peterson, P.A., Teyton, L., and Wilson, L.A. (1996). An  $\alpha\beta$  T cell receptor structure at 2.5 Å and its orientation in the TCR-MHC complex. *Science* **274**, 204-219.
- Garcia, K.C., Teyton, L., and Wilson, L.A. (1999). Structural basis of T cell recognition. *Annu. Rev. Immunol.* **17**, 369-397.
- Grant, K., Martin, C., Lassila, O., Suvilainen, R., Toivanen, A., and Toivanen, P. (1992). Immune capacity of the chicken bursaectomized at 60 hr of incubation; production of the immunoglobulins and specific antibodies. *Clin. Immunol. Immunopathol.* **23**, 459-469.
- Harlow, E., and Lane, D. (1988). *Antibodies: A Laboratory Manual* (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory).
- Hedrick, S.M., Engel, J., McEligott, D.L., Fink, P.J., Hsu, M.L., Hansburg, D., and Malis, L.A. (1988). Selection of amino acid sequences in the  $\beta$  chain of the T cell antigen receptor. *Science* **239**, 1541-1544.
- Hogenboom, H.R., and Winter, G. (1992). By-passing immunization. Human antibodies from synthetic repertoires of germline  $V_H$  gene segments rearranged in vivo. *J. Mol. Biol.* **227**, 381-388.
- Howard, J.G., Vicari, G., and Courtenay, B.M. (1979). Influence of molecular structure on the tolerogenicity of bacterial exantens. I. The  $\alpha$ -1-fused epitope of dextran B512. *Immunology* **29**, 585-597.
- Ishizuka, T., and Makioka, O. (1974). Inheritance of antibody specificity. I. Anti-4-hydroxy-5-bromo-3-nitrophenylacetyl of the mouse primary response. *J. Exp. Med.* **140**, 1495-1510.
- Imanishi, T., and Makioka, O. (1975). Inheritance of antibody specificity. II. Anti-4-hydroxy-5-bromo-3-nitrophenylacetyl of the mouse. *J. Exp. Med.* **141**, 840-854.
- Jalkanen, S., Grant, K., Jalkanen, M., and Toivanen, P. (1983). Immune capacity of the chicken bursaectomized at 60 hours of incubation: failure to produce immune, natural, and autoantibodies in spite of immunoglobulin production. *Cell. Immunol.* **60**, 363-373.
- Janeway, C.A., Jr., and Travers, P. (1997). *Immunobiology: The Immune System in Health and Disease*, Third Edition (New York: Current Biology Ltd./Garland Publishing Inc.), pp. 815-817.
- Jarin, J., and Chothia, C. (1999). The structure of protein-protein recognition sites. *J. Biol. Chem.* **274**, 16027-16030.
- Jourdain, S.H., Janse, E.M., van Rooijen, N., and Claassen, E. (1996). Inducible anti-polysaccharide antibody responses in the chicken. *Immunobiology* **193**, 385-395.
- Jorgensen, J.L., Esser, U., Farkas, de St. Gomb, B., Frey, P.A., and Davis, M.M. (1992). Mapping T-cell receptor-peptide contacts by variant peptide immunization of single-chain transgenics. *Nature* **355**, 224-230.
- Karntinen, M., Griffith, C.M., Markham, A.F., and Milstein, C. (1983). mRNA sequences define an unusually restricted IgG response to 2-phenylazoxalone and its early diversification. *Nature* **304**, 320-324.
- Kabat, E.A., Wu, T.T., and Biorolsky, H. (1977). Unusual distributions of amino acids in complementarity-determining (hypervariable) segments of heavy and light chains of immunoglobulins and their possible roles in specificity of antibody-combining sites. *J. Biol. Chem.* **252**, 6609-6616.
- Langman, R.E., and Cohn, M. (1987). The E-T (jaeger-haplo) paradox necessitates the concept of a unit of B-cell function: the protection. *Mol. Immunol.* **24**, 675-697.



- Li, H., Lebedeva, M.I., Ullera, A.S., Fields, B.A., Brunsen, M.B., and Manuzza, R.A. (1998). Structure of the Vd domain of a human  $\gamma$  T-cell antigen receptor. *Nature* 391, 502-508.
- López-Macías, C., Kalinka, U., Cascaho, M., Wabl, M., Hengartner, H., Zinkernagel, R.M., and Lamm, A. (1999). Secondary rearrangements and hypermutation generate sufficient B cell diversity to mount protective antiviral immunoglobulin responses. *J. Exp. Med.* 189, 1751-1759.
- Makela, O., Junt, M., and Becker, M. (1976). Inheritance of antibody specificity. II. A new  $V_H$  gene controls fine specificity of anti-azobenzene-arsenate complexed to the carbon atom 5 of hydroxyphenylacetic acid in mouse. *J. Exp. Med.* 143, 315-328.
- Makela, O., Kaartinen, M., Pelkonen, J., and Kaajala, K. (1978). Inheritance of antibody specificity. V. Anti-2-phenylazobenzene in the mouse. *J. Exp. Med.* 148, 1644-1650.
- Malmqvist, M. (1993). Dispersive interaction analysis using biosensor technology. *Nature* 367, 186-187.
- Marsner, T. (1990). Limits on heavy chain junctional diversity contribute to the recurrence of an antibody variable region. *Mol. Immunol.* 27, 503-511.
- Mason, D.W., and Williams, A.F. (1986). Kinetics of antibody reactions and the analysis of cell surface antigens. In *Handbook of Experimental Immunology*, Volume 1, D.M. Weir et al., eds. (Oxford: Blackwell Scientific Publications), pp. 381-387.
- McCormack, W.T., Tooleker, L.W., and Thompson, C.B. (1991). Avian B-cell development: generation of an immunoglobulin repertoire by gene conversion. *Annu. Rev. Immunol.* 9, 219-241.
- McKean, D., Huppi, K., Bell, M., Staudt, I., Gerhard, W., and Wüger, M. (1989). Generation of antibody diversity in the immune response of BALB/c mice to influenza virus hemagglutinin. *Proc. Natl. Acad. Sci. USA* 87, 3180-3184.
- Mian, I.S., Bradwell, A.R., and Olson, A.J. (1991). Structure, function and properties of antibody binding sites. *J. Mol. Biol.* 217, 133-151.
- Mushinski, E.B., and Potter, M. (1978). Idiotypes of galactin binding myeloma proteins and anti-galactin antibodies in mice. *J. Immunol.* 119, 1688-1693.
- Nissam, A., Hoogenboom, H.R., Tomlinson, I.M., Flynn, G., Midgley, C., Lane, D., and Winter, G. (1994). Antibody fragments from a "single pot" phage display library as immunochemical reagents. *EMBO J.* 13, 692-696.
- Noël, D., Bernardi, T., Navarro-Teulon, I., Marin, M., Martinetto, J.P., Duccioni, F., Mui, J.-C., Pau, B., Pechaczky, M., and Blard-Piccinini, M. (1996). Analysis of the individual contributions of immunoglobulin heavy and light chains to the binding of antigen using cell transfection and plasmon resonance analysis. *J. Immunol. Methods* 193, 177-187.
- Northrup, S.H., and Erickson, H.P. (1992). Kinetics of protein-protein association explained by Brownian dynamics computer simulation. *Proc. Natl. Acad. Sci. USA* 89, 3398-3402.
- Padian, E.A. (1980). On the nature of antibody combining sites: unusual structural features that may confer on these sites an enhanced capacity for binding ligands. *Proteins* 7, 112-124.
- Patten, P.A., Gray, N.S., Yang, P.L., Marks, C.B., Wiedemayer, G.J., Donifaci, J.J., Stevens, R.C., and Schultz, P.G. (1996). The immunological evolution of catalysis. *Science* 271, 1086-1091.
- Pauling, L. (1940). A theory of the structure and process of formation of antibodies. *J. Am. Chem. Soc.* 62, 2843-2857.
- Pavliak, I.L., Mushinski, E.B., Nisanci, A., and Potter, M. (1977). Evidence for the linkage of the IgCH locus to a gene controlling the idiotype specificity of anti-p-azobenzene-arsenate antibodies in strain A mice. *J. Exp. Med.* 137, 22-31.
- Potter, M., and Lieberman, R. (1970). Common individual antigenic determinants in five of eight BALB-c IgA myeloma proteins that bind phosphoryl choline. *J. Exp. Med.* 132, 737-751.
- Rajewsky, K., Förster, I., and Cumano, A. (1987). Evolutionary and somatic selection of the antibody repertoire in the mouse. *Science* 238, 1088-1094.
- Reinherz, E.L., Tan, K., Tang, L., Korn, P., Lu, J.-H., Xiong, Y., Hensley, J.E., Smolyar, A., Hale, B., Zhang, R., et al. (1999). The crystal structure of a T cell receptor in complex with peptide and MHC class II. *Science* 286, 1913-1921.
- Riblet, R., Blomberg, B., Weigert, M., Lieberman, R., Taylor, B.A., and Potter, M. (1979). Genetics of mouse antibodies. I. Linkage of the dextran response locus,  $D_{DEx}$ , to allotype. *Eur. J. Immunol.* 9, 773-777.
- Sant'Angelo, D.B., Waterbury, G., Preston-Hurlburt, P., Yoon, S.T., Medzhitov, R., Hong, S.-C., and Janeway, C.A., Jr. (1996). The specificity and orientation of a TCR to its peptide-MHC class II ligands. *Immunity* 4, 367-376.
- Seidl, K.J., MacKenzie, J.D., Wang, D., Kuntz, A.D., Kabat, E.A., and Herzberg, L.A. (1997). Frequent occurrence of identical heavy and light chain Ig rearrangements. *Int. Immunol.* 9, 689-702.
- Sher, A., and Colm, M. (1972). Inheritance of an idiotype associated with the immune response of inbred mice to phosphorylcholine. *Eur. J. Immunol.* 2, 319-326.
- Taylor, L.D., Carmack, C.E., Schwann, S.R., Mashayekh, R., Higgins, K.M., Kuo, C.C., Woodhouse, C., Kay, R.M., and Lemberg, N. (1992). A transgenic mouse that expresses a diversity of human sequence heavy and light chain immunoglobulins. *Nucleic Acids Res.* 20, 6287-6295.
- Taylor, L.D., Carmack, C.E., Huzar, D., Higgins, K.M., Mashayekh, R., Seagar, G., Schramm, S.R., Kuo, C.C., O'Donnell, S.L., Kay, R.M., et al. (1994). Human immunoglobulin transgenes undergo rearrangement, somatic mutation and class switching in mice that lack endogenous IgM. *Int. Immunol.* 6, 579-591.
- Tonegawa, S. (1983). Somatic generation of antibody diversity. *Nature* 302, 575-581.
- Wallick, S.C., Kabat, E.A., and Morrison, S.L. (1989). The effect of isotype and the J $\kappa$  region on antigen binding and idiotype expression by antibodies binding  $\alpha$ 1-6 dextran. *J. Immunol.* 142, 1235-1244.
- Wang, D.N., Chen, H.T., Liao, J., Akoliar, P.N., Sikder, S.K., Gruetz, F., and Kozet, E.A. (1996). Two families of monoclonal antibodies to  $\alpha$ 1-6 dextran, V $\alpha$ 19.1.2 and V $\alpha$ 2.14.7, show distinct patterns of J $\kappa$  and J $\kappa$  lineage usage and amino acid substitutions in CDH3. *J. Immunol.* 155, 3002-3010.
- Ward, E.S., Güssow, D., Göttsche, A.D., Jones, P.T., and Winter, G. (1989). Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli*. *Nature* 341, 544-546.
- Wiedenmayer, G.J., Patten, P.A., Wang, L.H., Schultz, P.G., and Stevens, R.C. (1997). Structural insights into the evolution of an antibody combining site. *Science* 276, 1655-1660.
- Well, J.C., and Raynaud, C.A. (1997). The chicken B cell compartment. *Science* 235, 1094-1098.
- Wells, J.A. (1996). Binding in the growth hormone receptor complex. *Proc. Natl. Acad. Sci. USA* 93, 1-6.